

日本特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

jc875 U.S. PTO
09/651058



08/30/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
this Office.

願年月日
Date of Application:

1999年 8月31日

願番号
Application Number:

平成11年特許願第246393号

願人
Applicant(s):

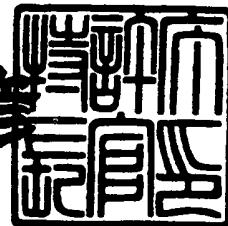
パイオニア株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 4月21日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



出証番号 出証特2000-3030183

【書類名】 特許願
【整理番号】 54P0080
【提出日】 平成11年 8月31日
【あて先】 特許庁長官殿
【国際特許分類】 G10L 7/08
G10L 9/18

【発明者】

【住所又は居所】 埼玉県川越市山田字西町25番地1 パイオニア株式会社
川越工場内

【氏名】 依田 正太郎

【特許出願人】

【識別番号】 000005016

【氏名又は名称】 パイオニア株式会社

【代理人】

【識別番号】 100063565

【弁理士】

【氏名又は名称】 小橋 信淳

【手数料の表示】

【予納台帳番号】 011659

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声認識システム

【特許請求の範囲】

【請求項1】 複数個の音検知手段と、

前記複数個の音検知手段から出力される各入力信号の中から音声認識に適した音声信号を決定する決定手段と、

前記決定手段で決定された音声信号に基づいて音声認識を行う音声認識手段とを具備することを特徴とする音声認識システム。

【請求項2】 前記決定手段は、前記複数個の音検知手段から出力される各入力信号の平均S/N値と平均音声パワーとを求め、前記平均S/N値と平均音声パワーがそれぞれ所定の閾値より大きい値となった入力信号を、前記音声認識に適した音声信号と決定することを特徴とする請求項1に記載の音声認識システム。

【請求項3】 前記決定手段は、前記平均S/N値と平均音声パワーがそれぞれ所定の閾値より大きな値となった音声信号のうち、平均S/N値と平均音声パワーの値の大きさに従って、音声認識に適した音声信号の候補順位を決定し、

前記音声認識手段は、前記候補順位の上位の音声信号に基づいて順次に音声認識を行うことを特徴とする請求項2に記載の音声認識システム。

【請求項4】 前記決定手段は、前記音声認識に適した音声信号以外の入力信号を雑音信号とすることを特徴する請求項1～3のいずれか1項に記載の音声認識システム。

【請求項5】 前記決定手段は、前記音声認識に適した音声信号以外の入力信号のうち、平均S/N値と平均音声パワーとが最も小さな値となる入力信号を雑音信号とすることを特徴とする請求項1～4のいずれか1項に記載の音声認識システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、発話音声によって電子機器の制御又は操作等を可能にする音声認識

システムに関するものである。

【0002】

【従来の技術】

従来、この種の音声認識システムとして、車載用オーディオシステムや車載用ナビゲーションシステム等の電子機器に適用されたものが知られている。

【0003】

音声認識システムを備えた車載用オーディオシステムでは、例えば、搭乗者が所望の放送局の局名を発話すると、音声認識システムが発話音声を音声認識し、認識結果に基づいてラジオ受信機の受信周波数を自動選局させることで、搭乗者の利便性の向上が図られている。

【0004】

また、車載用オーディオシステムに備えられているMD (Mini Disc) プレーヤーやCD (Compact Disc) プレーヤー等に、搭乗者が記録再生媒体を挿入し、その記録再生媒体に記録されている楽曲の名称を発話すると、音声認識システムが発話音声を音声認識して楽曲を自動的に再生させる等の機能も備えられている

【0005】

車載用ナビゲーションシステムでは、運転手等が行きたい目的地の名称を発話すると、音声認識システムが発話音声を音声認識して、現在地から目的地までの走行経路を地図表示させる等の機能が備えられ、この結果、運転手は運転操作に集中することができ安全運転の確保が図れる等の効果が得られるようになっている。

【0006】

【発明が解決しようとする課題】

ところで、上記従来の音声認識システムは、一人の発話者を対象とするものであったため、発話音声を入力するためのマイクロフォンが使用頻度の高い運転手に最も近い位置に1個だけ設けられていた。

【0007】

このため、例えば、マイクロフォンから遠い位置に着座している他の搭乗者は

、マイクロフォンの音声入力レベルを十分に確保すべく、マイクロフォンに向かって大声で発話しなければならないという問題があった。また、音声認識システムの音声認識率を高めるためには、他の搭乗者は発話音声を車内雑音の影響を受けることなくマイクロフォンに入力させるべく、マイクロフォンに向かって大声で発話しなければならないという問題があった。

【0008】

本発明はこのような問題点に鑑みて成されたものであり、より利便性の高い音声認識システムを提供することを目的とする。

【0009】

【課題を解決するための手段】

上記目的を達成するための本発明は、複数個の音検知手段と、上記複数個の音検知手段から出力される各入力信号の中から音声認識に適した音声信号を決定する決定手段と、上記決定手段で決定された音声信号に基づいて音声認識を行う音声認識手段とを具備することを特徴とする。

【0010】

また、上記決定手段は、上記複数個の音検知手段から出力される各入力信号の平均S/N値と平均音声パワーとを求め、上記平均S/N値と平均音声パワーがそれぞれ所定の閾値より大きい値となった入力信号を上記音声認識に適した音声信号と決定することを特徴とする。

【0011】

上記決定手段は、上記平均S/N値と平均音声パワーがそれぞれ所定の閾値より大きな値となった音声信号のうち、平均S/N値と平均音声パワーの値の大きさに従って音声認識に適した音声信号の候補順位を決定し、上記音声認識手段は、上記候補順位の上位の音声信号に基づいて順次に音声認識を行うことを特徴とする。

【0012】

また、上記決定手段は、上記音声認識に適した音声信号以外の入力信号を雑音信号とすることを特徴する。

【0013】

また、上記決定手段は、上記音声認識に適した音声信号以外の入力信号のうち、平均S/N値と平均音声パワーとが最も小さな値となる入力信号を雑音信号とすることを特徴とする。

【0014】

これらの構成によれば、発話者が所望の発話をすると、複数個の音検知手段から出力される各入力信号の中から、音声認識を行うのに適した音声信号と雑音信号が自動的に決定され、その決定された音声信号と雑音信号に基づいて音声認識が行われる。この結果、発話者は特定の音検知手段に対して発話するという意識を持つことなく、発話するだけで適切な音声認識がなされることとなり、利便性の向上等が実現される。

【0015】

【発明の実施の形態】

以下、本発明の実施の形態を図面を参照して説明する。尚、一実施形態として、車載用オーディオシステムや車載用ナビゲーションシステム等の車載用電子機器を発話音声によって制御又は操作することを可能にする音声認識システムについて説明する。

【0016】

図1は、本実施形態に係る音声認識システムの構成を示すブロック図である。同図において、本音声認識システムは、音検知手段としての複数個のマイクロフォンM₁～M_N、複数の前置回路CC₁～CC_N、マルチプレクサ1、A/D変換器2、デマルチプレクサ3、記憶部4、発話音声検出部5、データ解析部6、音声認識部7、制御部8、発話スイッチ9を備えて構成されている。

【0017】

ここで、複数の前置回路CC₁～CC_N、マルチプレクサ1、A/D変換器2、デマルチプレクサ3、記憶部4、発話音声検出部5、データ解析部6、制御部8、は、音声認識を行うのに適した音声信号と雑音信号とを決定するための決定手段を構成している。

【0018】

発話スイッチ9は1個だけ備えられ、運転手の座席近傍、例えば車両のフロントダッシュボードや前部ドアの一端に設けられる。

【0019】

制御部8は、マイクロプロセッサ(MPU)を有し、発話スイッチ9がオン操作されてオン信号SWが供給されると、マイクロフォンM₁～M_Nに集音動作を開始させると共に、本音声認識システム全体の動作を制御する。尚、詳細については動作説明と共に述べるが、発話音声検出部5には、どのマイクロフォンに発話がなされたかを判断するための連続回数カウンタFC₁～FC_Nが備えられている。

【0020】

各マイクロフォンM₁～M_Nは、各搭乗者の発話音声を集音し易い位置、例えば各搭乗者の座席近傍に設けられる。

【0021】

一例として4個のマイクロフォンM₁～M₄を車室内に設置する場合には、図2(a)の平面図に示すように、マイクロフォンM₁, M₂を運転席と助手席の前方にそれぞれ設置し、マイクロフォンM₃, M₄を後部座席の前方、例えば天井部分や、運転席と助手席の背部等に設置することで、各マイクロフォンM₁～M₄と各搭乗者とを対応付けるようとする。

【0022】

また、他の一例として、図2(b)の平面図に示すように、マイクロフォンM₁, M₂を運転席側の前部ドアと助手席側の前部ドアにそれぞれ設置し、マイクロフォンM₃, M₄を各後部座席側の後部ドア等に設置することにより、各マイクロフォンM₁～M₄と各搭乗者とを対応付けるようとする。

【0023】

また、更に他の一例として、図2(a)(b)を組み合わせた位置に各マイクロフォンM₁～M₄を設置してもよい。つまり、マイクロフォンM₁を図2(a)に示す運転席の前方に設置する場合には図2(b)に示す運転席側の前部ドアには設置せず、マイクロフォンM₁を図2(b)に示す運転席側の前部ドアに設置

する場合には図2(a)に示す運転席の前方には設置しないように、図2(a)(b)に示す何れか一方の位置を選択して、運転席に着座する運転手に対して1個のマイクロフォンM₁を設置するようとする。他のマイクロフォンM₂～M₄についても同様に、図2(a)(b)に示す何れか一方の位置を選択して、それぞれ1個ずつ設置する。

【0024】

また、座席数の多いワゴン車等の場合には、図3(a)(b)の平面図に示すように、座席数に応じてより多くのマイクロフォンM₁～M₆等を設置し、これらのマイクロフォンM₁～M₆も各搭乗者の発話音声を集音し易い位置に1個ずつ設ける。また、図3(a)(b)を組み合わせた位置に各マイクロフォンM₁～M₄を1個ずつ設置してもよい。

【0025】

尚、上述したマイクロフォンの各配置例はあくまでも一例である。実際には、本発明のシステムで使用するシステム情報は、予め搭乗者からマイクロフォンまでの音響伝達特性を考慮して構築される。そのため、マイクロフォンの設置条件が厳密に制限されることはない。また、マイクロフォンの数は、車種等に応じて予め決められている最大搭乗者数と同数又はそれより少ない数に決めることができる。

【0026】

また、単に各マイクロフォンから各搭乗者までの距離を等しくするというではなく、車室内の音響特性を実験等によって予め解析しておき、この解析結果に基づいて、各マイクロフォンから各搭乗者までのそれぞれの音響伝達特性がほぼ等しくなるように、それぞれの距離と設置場所が決めることができる。

【0027】

再び図1において、各マイクロフォンM₁～M_Nには、前置回路C C₁～C C_Nが対応付けて接続され、これによってNチャンネル分の信号処理系統が構成されている。

【0028】

各前置回路C C₁～C C_Nには、各マイクロフォンM₁～M_Nから供給される入力

信号 $S_1 \sim S_N$ を信号処理可能な振幅レベルに増幅する増幅器（図示省略）と、増幅器で増幅された入力信号のうち所定周波数成分のみを通過させるバンドパスフィルタ（図示省略）が備えられており、各バンドパスフィルタを通過した入力信号 $S'_1 \sim S'_N$ をマルチプレクサ1に供給するようになっている。

【0029】

上記の各バンドパスフィルタは、入力信号 $S_1 \sim S_N$ に含まれる低域雑音を除去するための低域カットオフ周波数 f_L （例えば、 $f_L = 100\text{ Hz}$ ）と、ナイキスト周波数を考慮した高域カットオフ周波数 f_H が設定されており、これら低域カットオフ周波数 f_L と高域カットオフ周波数 f_H との周波数範囲内に、人間が発話する音声の音声周波数帯域が含まれるように設定されている。

【0030】

マルチプレクサ1は、図4に示すように、Nチャンネル分のアナログスイッチ $A S_1 \sim A S_N$ を備えて構成されている。各アナログスイッチ $A S_1 \sim A S_N$ の入力端に各前置回路 $CC_1 \sim CC_N$ からの入力信号 $S'_1 \sim S'_N$ が供給され、各アナログスイッチ $A S_1 \sim A S_N$ の出力端は共通接続されてA/D変換器2に接続されている。そして、制御部8から供給されるチャンネル切換え信号 $CH_1 \sim CH_N$ に応じて、入力信号 $S'_1 \sim S'_N$ を排他的に切換えてA/D変換器2に供給する。

【0031】

A/D変換器2は、マルチプレクサ1から順次に供給される入力信号 $S'_1 \sim S'_N$ を所定のサンプリング周波数 f に同期してそれぞれデジタルの入力データ $D_1 \sim D_N$ に変換し、デマルチプレクサ3に供給する。

【0032】

上記のサンプリング周波数 f は、制御部8からのサンプリングクロック CK_{AD} によって設定されると共に、アンチエイリアシングを考慮した周波数に決められている。より具体的には、サンプリング周波数 f は、前置回路 $CC_1 \sim CC_N$ に備えられている上記バンドパスフィルタの高域カットオフ周波数 f_H の約2倍以上に決められ、例えば、8 kHzないし11 kHzの範囲内の周波数に決められている。

【0033】

デマルチプレクサ3は、図4に示すように、Nチャンネル分のアナログスイッチAW₁～AW_Nを備えて構成されている。各アナログスイッチAW₁～AW_Nの入力端は共通接続されてA/D変換器2の出力端に接続され、各アナログスイッチAW₁～AW_Nの各出力端は、記憶部4に備えられているNチャンネル分の記憶領域ME₁～ME_Nに接続されている。そして、制御部8から供給されるチャンネル切換え信号CH₁～CH_Nに応じて、入力データD₁～D_Nを排他的に切換えて各記憶領域ME₁～ME_Nに供給する。

【0034】

ここで、図5のタイミングチャートを参照して、マルチプレクサ1とA/D変換器2及びデマルチプレクサ3の動作を説明する。同図において、発話スイッチ9がオン操作されると、それによって生じるオン信号SWを制御部8が受信し、サンプリングクロックCK_{ADC}とチャンネル切換え信号CH₁～CH_Nを出力する。

【0035】

サンプリングクロックCK_{ADC}は、サンプリング周波数fの逆数1/fの周期(サンプリング周期)Tの間にN回の論理反転を繰り返すパルス波形であり、各チャンネル切換え信号CH₁～CH_Nは、サンプリングクロックCK_{ADC}の1周期T/N毎に順番に論理“1”となるパルス波形である。

【0036】

マルチプレクサ1は、チャンネル切換え信号CH₁～CH_Nが順番に論理“1”となる周期T/Nに同期して排他的に導通／非導通の切換え動作を行う。このため、入力信号S₁'～S_N'は、周期T/Nに同期して順番にA/D変換器2に供給されてデジタルデータD₁～D_Nに変換される。一方、デマルチプレクサ3も、チャンネル切換え信号CH₁～CH_Nが順番に論理“1”となる周期T/Nに同期して排他のに導通／非導通の切換え動作を行う。このため、A/D変換器2から出力される入力データD₁～D_Nは、周期T/Nに同期して記憶領域ME₁～ME_Nに振り分けられて記憶される。

【0037】

このように、サンプリング周期T ($= 1/f$) の間にNチャンネル分の入力信号 $S_1' \sim S_N'$ をサンプリングしてこれを繰り返すことで、1個のA/D変換器2でも、Nチャンネル分の入力データ $D_1 \sim D_N$ をサンプリング周波数fに同期して生成し、所定の記憶領域 $ME_1 \sim ME_N$ に振り分けて記憶させることが可能となっている。

【0038】

記憶部4は、半導体メモリで形成され、上述したNチャンネル分の記憶領域 $ME_1 \sim ME_N$ を備えている。すなわち、各記憶領域 $ME_1 \sim ME_N$ は、各マイクロフォン $M_1 \sim M_N$ に対応付けて備えられている。

【0039】

更に図4に示すように、各記憶領域 $ME_1 \sim ME_N$ には、それぞれの入力データ $D_1 \sim D_N$ を所定サンプル数のフレーム単位で記憶するために、複数のフレーム領域 $MF_1, MF_2 \dots$ が備えられている。

【0040】

記憶領域 ME_1 を代表して述べれば、各フレーム領域 $MF_1, MF_2 \dots$ は、デマルチプレクサ3から供給される入力データ D_1 を、制御部8からのアドレス信号ADR₁に従って予め決められたサンプル数（本実施形態では、256サンプル）ずつ順番に記憶するようになっている。つまり、図5に示した期間（256×T）をフレーム期間TFと呼び、この1フレーム期間（1TF）毎に、各フレーム領域 $MF_1, MF_2 \dots$ に256個ずつの入力データ D_1 を記憶させるようになっている。また、各フレーム領域 $MF_1, MF_2 \dots$ に記憶される1フレーム期間（1TF）分の入力データをフレームデータと呼んでいる。

【0041】

そして、残余の記憶領域 $ME_2 \sim ME_N$ に備えられている各フレーム領域 $MF_1, MF_2 \dots$ にも、それぞれの入力データ $D_2 \sim D_N$ を1フレーム期間（1TF）毎に、256サンプルずつを記憶させるようになっている。

【0042】

発話音声検出部5とデータ解析部6は、DSP (Digital Signal Processor)

で形成されている。

【0043】

発話音声検出部5は、各記憶領域 $M_E_1 \sim M_E_N$ の各フレーム領域 $M_F_1, M_F_2 \dots$ にフレームデータが記憶される度に、最新のフレームデータのLPC残差を演算し、その演算値が所定の閾値THD1以上となるか判定する。演算値が所定の閾値THD1以上となった場合には、最新のフレームデータを発話によって生じた音声フレームデータであると判定し、演算値が所定の閾値THD1未満の場合には、最新のフレームデータを発話によらない入力データ、すなわち車室内の雑音によって生じた雑音フレームデータであると判定する。

【0044】

更に発話音声検出部5は、LPC残差の演算値が3フレーム期間(3TF)に亘って連続して閾値THD1以上となった場合に、その3フレーム期間(3TF)のフレームデータを明らかに発話によって生じた音声フレームデータであると確定し、その確定結果を示す音声検出データDCT1を制御部8へ転送する。

【0045】

より詳細には、各記憶領域 $M_E_1 \sim M_E_N$ の各フレーム領域 $M_F_1, M_F_2 \dots$ に記憶されるフレームデータのLPC残差を各チャンネル毎に個別演算し、更に、各チャンネル毎に求められたLPC残差の演算値を閾値THD1と比較することで、各チャンネル毎に上記の発話による音声フレームデータか否かの判定を行うようになっている。

【0046】

つまり、マイクロフォン M_1 に関連する第1チャンネルのLPC残差の演算値が ε_1 、マイクロフォン M_2 に関連する第2チャンネルのLPC残差の演算値が ε_2 、以下同様にマイクロフォン $M_3 \sim M_N$ に関連する第3～第Nチャンネルの各LPC残差の演算値が $\varepsilon_3 \sim \varepsilon_N$ であった場合には、各々の演算値 $\varepsilon_1 \sim \varepsilon_N$ を閾値THD1と比較する。そして、閾値THD1以上となったチャンネルに該当するフレームデータを発話によって生じた音声フレームデータであると判定し、更に、3フレーム期間(3TF)に亘って連続してLPC残差の演算値が閾値THD1以上となったチャンネルの音声フレームデータを、明らかに発話によって生じた

音声フレームデータであると確定することとしている。

【0047】

例えばマイクロフォンM₁に向かって発話が行われ、残余のマイクロフォンM₂～M_Nには発話音声が入力されなかったような場合には、マイクロフォンM₁に関連するチャンネルの記憶領域M E₁に記憶されたフレームデータだけが発話によって生じた音声フレームデータであると判定及び確定され、残余のマイクロフォンM₂～M_Nに関連するチャンネルの記憶領域M E₂～M E_Nに記憶されたフレームデータは、車室内の雑音によって生じた雑音フレームデータであると判定される。

【0048】

また、例えばマイクロフォンM₁に向かって発話が行われると共に、マイクロフォンM₂にはその発話音声が回り込んで入力され、残余のマイクロフォンM₃～M_Nには発話音声が入力されなかったような場合には、マイクロフォンM₁とM₂に関連するチャンネルの記憶領域M E₁とM E₂に記憶された両方のフレームデータが発話によって生じた音声フレームデータであると判定及び確定され、残余のマイクロフォンM₃～M_Nに関連するチャンネルの記憶領域M E₃～M E_Nに記憶されたフレームデータは、雑音フレームデータであると判定される。

【0049】

このように、発話音声検出部5は、記憶領域M E₁～M E_Nに記憶されたフレームデータの各LPC残差を演算して閾値THD1と比較することで、何れのマイクロフォンに発話音声が入力されたかの判定と、発話音声が入力されたフレーム期間の判定を行い、これらの判定情報を有する音声検出データDCT1を制御部8へ転送する。

【0050】

尚、音声検出データDCT1は、上記の3フレーム期間以上連續して音声フレームデータが記憶された記憶領域（以下、音声記憶チャンネルという）とそのフレーム領域（以下、音声記憶フレームという）とを示す所定のコードデータとして制御部8へ転送される。

【0051】

具体例として、音声検出データDCT1は、 $DCT1 \{CH_1 (TF_1, TF_2 \sim TF_m), CH_2 (TF_1, TF_2 \sim TF_m), \dots, CH_N (TF_1, TF_2 \sim TF_m) \}$ 等の一般形式のデータ構造を有している。 $CH_1 \sim CH_N$ は各チャンネルを表すフラグデータ、 $TF_1, TF_2 \sim TF_m$ は、各フレーム領域 $MF_1, MF_2 \sim MF_m$ に対応するフラグデータとなっている。

【0052】

そして、例えばマイクロフォン M_1 のみに発話音声が入力され、且つ第3番目以降のフレーム領域 $MF_3, MF_4 \dots$ に音声フレームデータが記憶されたされたような場合は、 $DCT1 \{1(0,0,1,1,\sim1), 0(0,0,0\sim0), \dots, 0(0,0,0\sim0)\}$ となるバイナリーコードの音声検出データDCTが制御部8へ転送される。

【0053】

この音声検出データDCT1が制御部8へ転送されると、制御部8は音声検出データDCT1に基づいて、音声記憶チャンネルと音声記憶フレームを示す制御データCNT1を生成して、データ解析部6に供給する。

【0054】

データ解析部6は、最適音声決定部6a、雑音決定部6b、平均S/N値演算部6c、平均音声パワー演算部6d、平均雑音パワー演算部6e、音声条件テーブル6f、雑音選択テーブル6gを備えて構成されており、制御部8からの制御データCNT1が供給されると、音声認識に適した音声フレームデータと雑音フレームデータを決定するための処理を開始する。

【0055】

ここで、平均音声パワー演算部6dは、制御データCNT1から音声記憶チャンネル及び音声記憶フレームの情報を取得し、記憶領域 $ME_1 \sim ME_N$ のうちのこれら音声記憶チャンネル及び音声記憶フレームに該当する記憶領域から音声フレームデータを読み出して、各チャンネル毎に、音声フレームデータの平均音声パワーP(n)を演算する。尚、平均音声パワーP(n)の変数nはチャンネル番号を示す変数である。

【0056】

例えば、図6 (a) ~ (d) に示すように、チャンネルCH₁~CH₄に当たる記憶領域ME₁~ME₄に音声フレームデータが記憶されている場合には、各チャンネル毎に、発話が開始された時点t_sから予め決められた複数フレーム期間(m₂×TF)内に該当する複数の音声フレームデータの平均音声パワーP(1)~P(4)を演算する。尚、平均音声パワーP(n)は、フレーム期間(m₂×TF)内の音声フレームデータを2乗加してフレーム期間(m₂×TF)で除算することにより演算される。

【0057】

平均雑音パワー演算部6eは、制御データCNT1から音声記憶チャンネル及び音声記憶フレームの情報を取得し、記憶領域ME₁~ME_Nのうちのこれら音声記憶チャンネルに該当する記憶領域から、音声フレームデータよりも複数フレーム期間(m₁×TF)前の雑音フレームデータを読み出し、各チャンネル毎に、雑音フレームデータの平均雑音パワーNP(n)を演算する。尚、NP(n)の変数nは音声チャンネルを示す変数であり、また、平均雑音パワーNP(n)は、フレーム期間(m₁×TF)内の雑音フレームデータを2乗加算してフレーム期間(m₁×TF)で除算することにより演算される。

【0058】

例えば、図6 (a) ~ (d) に示すように、チャンネルCH₁~CH₄に当たる記憶領域ME₁~ME₄に音声フレームデータが記憶されている場合には、発話が開始された時点t_s(音声フレームデータの記憶開始時点)よりも複数フレーム期間(m₁×TF)前に該当する複数の雑音フレームデータの平均雑音パワーNP(n)が演算される。

【0059】

平均S/N値演算部6cは、平均音声パワー演算部6dで演算された平均音声パワーP(n)と平均雑音パワー演算部6eで演算された平均雑音パワーNP(n)に基づいて、音声チャンネル毎に、信号対雑音比の値を表すS/N値SN(n)を演算する。

【0060】

例えば、図6 (a) ~ (d) に示すように、チャンネルCH₁~CH₄が音声チャンネルであった場合には、次式(1)~(4)に基づいて、各チャンネルCH₁~CH₄のS/N値SN(1)~SN(4)を演算する。

【0061】

$$SN(1) = P(1) / NP(1) \quad \dots (1)$$

$$SN(2) = P(2) / NP(2) \quad \dots (2)$$

$$SN(3) = P(3) / NP(3) \quad \dots (3)$$

$$SN(4) = P(4) / NP(4) \quad \dots (4)$$

また、上記式(1)~(4)によって演算されたS/N値SN(1)~SN(4)のそれぞれの対数値を、各チャンネルCH₁~CH₄のS/N値SN(1)~SN(4)としてもよい。

【0062】

最適音声決定部6aは、平均S/N値演算部6cで求められた平均S/N値SN(n)を所定の閾値THD2と比較すると共に、平均音声パワー演算部6dで求められた平均音声パワーP(n)を所定の閾値THD3と比較し、これらの比較結果を図7に示す音声条件テーブル6fと対比することで、何れのチャンネルの音声フレームデータが音声認識処理に適用するのに好適か判定する。

【0063】

つまり、音声条件テーブル6fには、図7に示すように、平均S/N値と閾値THD2との関係と平均音声パワーと閾値THD3との関係とによって音声フレームデータを等級(ranking)付けるための参照データが記憶されており、最適音声決定部6aは、上記の比較結果に基づいて音声条件テーブル6fを参照することで、音声認識処理に適した音声フレームデータを等級付けし、最も上位の等級の音声フレームデータを音声認識処理に最も適していると判定する。

【0064】

具体的には、平均S/N値が閾値THD2以上で且つ平均パワー値が閾値THD3以上となった音声フレームデータを等級1(Rnk1)、平均S/N値が閾値THD2以上で且つ平均パワー値が閾値THD3未満となった音声フレームデータ

を等級2 (Rnk2) 、平均S/N値が閾値THD2未満で且つ平均パワー値が閾値THD3以上となった音声フレームデータを等級3 (Rnk3) 、平均S/N値が閾値THD2未満で且つ平均パワー値が閾値THD3未満となった音声フレームデータを等級4 (Rnk4) と判定する。

【0065】

更に、全てのチャンネルの音声フレームデータの中で、平均S/N値と平均パワー値が最大となった音声フレームデータを等級0 (Rnk0) と判定する。

【0066】

そして、等級0 (Rnk0) となった音声フレームデータを音声認識処理に最も適した候補（第1候補）と判定する。更に、等級1 (Rnk1) となった音声フレームデータを音声認識処理に適した次の候補（第2候補）と判定する。更にまた、等級1 (Rnk1) となったチャンネルが複数存在した場合には、平均S/N値及び平均パワー値の大きなものほど上位の候補と判定する。

【0067】

更に、等級2 (Rnk2) ~等級4 (Rnk4) に該当する音声フレームデータを音声認識処理に適さないものとして音声認識処理の対象外とする。

【0068】

このように、最適音声決定部6aは、平均S/N値SN(n)と平均音声パワーP(n)をそれぞれ閾値THD2, THD3と比較し、これらの比較結果を図7に示す音声条件テーブル6fと対比することで、音声認識処理に適した音声フレームデータを判定し、更に、音声認識処理に適した音声フレームデータに優先順位を付ける。そして、優先順位を示す音声候補データDCT2を制御部8へ転送する。

【0069】

雑音決定部6bは、最適音声決定部6aで求められたNチャンネル分の全等級の組み合わせを図8に示す雑音選択テーブル6gと対比し、組み合わせが一致した場合のチャンネルを雑音のチャンネルと判定する。

【0070】

例えば、図8中の「場合1」に示すように、最適音声決定部6aで求められた

第1チャンネルCH₁以降の各チャンネルの等級が(Rnk0), (Rnk1), (Rnk2), (Rnk3), ……であった場合には、雑音のチャンネルを第3チャンネルCH₃と判定する。そして、雑音のチャンネルを示す雑音候補データDCT3を制御部8へ転送する。

【0071】

したがって、最適音声決定部6aで音声認識に適した音声フレームデータの候補が判定されると、雑音決定部6bは、図8中の各「場合」を参照することにより、その音声認識に適した音声フレームデータの候補に対応する雑音チャンネルを判定する。この結果、音声認識に適した音声フレームデータの候補と、雑音を集音したマイクロフォンによって得られた雑音データとが関連付けて判定される。

【0072】

尚、図8に示す雑音選択テーブル6gの各場合1, 2, 3…は、全てのマイクロフォンM₁～M_Nを車室内に実際に設置して、実際に搭乗者が様々な位置で発話したときに行われる音響特性の実験結果に基づいて予め設定されている。

【0073】

このように音声候補データDCT2と雑音候補データDCT3が制御部8に供給されると、制御部8は、記憶領域ME₁～ME_Nのうち、音声候補データDCT2に基づいて第1候補のチャンネルに該当する記憶領域をメモリアクセスし、最も音声認識処理に適した音声フレームデータを読み出して音声認識部7へ供給する。更に、雑音候補データDCT3に基づいて雑音のチャンネルに該当する記憶領域をメモリアクセスし、雑音フレームデータを読み出して音声認識部7へ供給する。

【0074】

音声認識部7は、記憶部4により供給される音声フレームデータと雑音フレームデータに基づいてSS(Spectrum Subtraction)処理、エコーキャンセラー、ノイズキャンセラー、CMN処理等の周知の処理を行うことにより、音声フレームデータ中に含まれている雑音成分を除去し、その雑音成分の除去された音声フレームデータに基づいて音声認識処理を行い、更に、音声認識結果のデータDou

tを出力する。

【0075】

尚、音声認識部7が上記の音声認識処理に適したフレームデータと雑音のフレームデータに基づいて音声認識処理を行った結果、適切な音声認識結果が得られなかった場合には、制御部8は、音声認識処理に適した次候補チャンネルに該当する記憶領域をメモリアクセスしてその音声フレームデータを音声認識部7へ転送する。そして、以下同様に適切な音声認識結果が得られるまで、制御部8は順番に次候補チャンネルの音声フレームデータを音声認識部7に供給する。

【0076】

次に、かかる構成を有する本音声認識システムの動作例を図9及び図10に示すフローチャートに基づいて説明する。尚、図9は、マイクロフォンM₁～M_Nで集音し入力データD₁～D_Nをフレームデータとして記憶部4に記憶するまでの動作、図10は、データ解析部6が最適の音声フレームデータと雑音データを決定する際の動作を示している。

【0077】

図9において、ステップ100では発話スイッチ9がオン操作されるまで待機し、発話スイッチ9がオン操作されるとステップ102に移行して初期化処理が行われる。この初期化処理では、制御部8内に設けられているチャネル番号カウンタの計数値nと、フレーム番号カウンタの計数値mと、連続回数カウンタFC₁～FC_Nの全ての値F(1)～F(N)をクリアする。

【0078】

尚、上記のチャネル番号カウンタは、マイクロフォンM₁～M_Nの各チャンネルを計数値nによって指定するために設けられている。上記のフレーム番号カウンタは、記憶領域ME₁～ME_Nに備えられている各フレーム領域MF₁、MF₂、MF₃…の番号(アドレス)を計数値mによって指定するために設けられている。

【0079】

上記の連続回数カウンタはFC₁～FC_Nは、各チャンネルに対応付けてN個備えられている。すなわち、第1の連続回数カウンタFC₁は第1チャンネル、第2の連続回数カウンタFC₂は第2チャンネル、以下同様にして第Nの連続回数

カウンタ FC_N は第 N チャンネルに対応付けて設けられている。そして、閾値 THD_1 より大きなLPC残差 ε_n が連続して3フレーム以上続いたか否かと、その3フレーム以上続いたチャンネルを判定するために設けられている。更に、その3フレーム以上続いたチャンネルを発話入力のなされたチャンネルと判定するために設けられている。

【0080】

次に、ステップ104において、記憶領域 $ME_1 \sim ME_N$ のそれぞれの第1番目のフレーム領域 MF_1 を設定する。すなわち、フレーム領域の番号 m を $m = 1$ に設定する。

【0081】

次に、ステップ106と108において、マイクロフォン $M_1 \sim M_N$ に集音動作を開始させ、それによって得られる入力データ $D_1 \sim D_N$ を記憶領域 $ME_1 \sim ME_N$ の第1番目の各フレーム領域 MF_1 に1フレーム分ずつ記憶させる。

【0082】

上記の1フレーム分ずつの入力データ $D_1 \sim D_N$ が記憶されると、次にステップ110において、第1 ($n = 1$) のチャンネルに該当する記憶領域 ME_1 を指定した後、ステップ112において記憶領域 ME_1 の第1番目 ($m = 1$) のフレーム領域 MF_1 に記憶されたフレームデータのLPC残差 ε_n ($n = 1$) を演算する。

【0083】

次に、ステップ114において、そのLPC残差 ε_n と閾値 THD_1 とを比較する。 $\varepsilon_n \geq THD_1$ のときにはステップ116に移行して第1チャンネルに該当する連続回数カウンタ FC_1 の値 $F(1)$ をインクリメント (1を加算) する。 $\varepsilon_n < THD_1$ のときにはステップ118に移行して連続回数カウンタ FC_1 の値 $F(1)$ をクリアする。

【0084】

したがって、 $\varepsilon_n \geq THD_1$ となると、連続回数カウンタ FC_1 の値 $F(1)$ は1になるので、この値 $F(1)$ は、第1チャンネルのマイクロフォン M_1 に対し1フレーム分の発話入力がなされたことを示すことになる。

【0085】

一方、 $\varepsilon_n < THD_1$ となると、連続回数カウンタ FC_1 の値 $F(1)$ が 0 にクリアされるので、この値 $F(1)$ は、第 1 チャンネルのマイクロフォン M_1 に対し発話入力がなされていないことを示すことになる。

【0086】

次に、ステップ 120において $n = N$ か調べることにより、全てのチャンネルにおける LPC 残差 ε_n を演算したか判断する。 $n = N$ でなければステップ 122 に移行して、 $n = n + 1$ とすることにより、次のチャンネルを設定してステップ 112 からの処理を繰り返す。すなわち、ステップ 112～122 の処理を N 回繰り返すことで、記憶領域 $ME_1 \sim ME_N$ の各フレーム領域 MF_1 に記憶されたフレームデータの LPC 残差 ε_n と閾値 THD_1 とを比較する。そして、LPC 残差 ε_n が閾値 THD_1 以上となったときには、そのチャンネル番号 n に該当する連続回数カウンタ FC_n の値 $F(n)$ をインクリメントする。

【0087】

次に、上記ステップ 120において $n = N$ となると、全チャンネルについての処理が完了したと判断してステップ 124 に移行する。

【0088】

ステップ 124 では、連続回数カウンタ $FC_1 \sim FC_N$ の値 $F(1) \sim F(N)$ のうち、値が 3 以上となったものがあるか判断する。値が 3 以上となったものが無ければ、すなわち、何れの値 $F(1) \sim F(N)$ も 2 以下であればステップ 126 に移行する。

【0089】

ステップ 126 では、 $m = m + 1$ とすることにより記憶領域 $ME_1 \sim ME_N$ の第 2 番目の各フレーム領域 MF_2 を設定する。そして、ステップ 106～124 の処理が繰り返される。

【0090】

これにより、各フレーム領域 MF_2 に入力データを記憶させ（ステップ 106, 108）、更に各フレーム領域 MF_2 に記憶された各フレームデータの LPC 残差 ε_2 を閾値 THD_1 と比較し（ステップ 110～114）、その比較結果に

基づいて各連続回数カウンタ $FC_1 \sim FC_N$ の各値 $F(1) \sim F(N)$ をインクリメント又はクリアすることになる。

【0091】

そして、再びステップ124において、連続回数カウンタ $FC_1 \sim FC_N$ の値 $F(1) \sim F(N)$ のうち、値が3以上となったものがあるか判断し、無ければステップ130に移行して、更に $m = m + 1$ とすることにより記憶領域 $ME_1 \sim ME_N$ の次の各フレーム領域 MF_3 を設定してステップ106～124の処理を繰り返す。

【0092】

こうしてステップ106～126の処理を繰り返し、ステップ124において、連続回数カウンタ $FC_1 \sim FC_N$ の値 $F(1) \sim F(N)$ のうち、少なくとも1つの連続回数カウンタの値が3以上となると、ステップ128の処理に移行する。

【0093】

つまり、ステップ124では、連続回数カウンタ $FC_1 \sim FC_N$ の値 $F(1) \sim F(N)$ を調べ、閾値 THD1 より大きなLPC残差 ϵ_n が連続して3フレーム以上続いた場合に限り、そのチャンネルに該当する記憶領域に記憶されたフレームデータを音声フレームデータであると判断し決定する。

【0094】

次に、ステップ128では、上記の閾値 THD1 より大きなLPC残差 ϵ_n が連続して3フレーム以上続いたと判断された連続解するカウンタの値が更に5に達したか否か判断し、値が5に達していないければステップ126の処理を行って、再びステップ106～128の処理を繰り返す。

【0095】

すなわち、或るチャンネル n に該当する連続回数カウンタ FC_n の値 $F(n)$ が3になったときに、残余のチャンネルに該当する連続回数カウンタの値が1又は2になっている場合がある。こうした場合には、残余のチャンネルに該当する記憶領域に記憶されたフレームデータも音声フレームデータである可能性がある。

【0096】

そこで、上記ステップ106～128の処理を更に2回繰り返すことにより、残余のチャンネルに該当する記憶領域に記憶されたフレームデータが音声フレー

ムであるか否か確認することとしている。

【0097】

次に、ステップ128の判断が「YES」となると、ステップ130に移行し、音声フレームデータが記憶された記憶領域と雑音フレームデータが記憶された記憶領域との情報を有する音声検出データDCT1を制御部8へ転送した後、図10に示す処理に移行する。

【0098】

図10の処理に移行すると、まずステップ200において、各チャンネルの平均音声パワーP(n)と平均雑音パワーNP(n)及び平均S/N値SN(n)を演算する。次に、ステップ202において、図7に示した音声条件テーブル6fに基づいて音声認識に適した音声フレームデータの候補を決定する。更に、ステップ204において、図8に示した雑音選択テーブル6gに基づいて音声認識に適した雑音フレームデータを決定する。

【0099】

そして、ステップ206において、データ解析部6から制御部8へ、音声認識に適した音声フレームデータの候補を示す音声候補データDCT2と雑音フレームデータを示す雑音候補データDCT3が転送される。つまり、音声候補データDCT2と雑音候補データDCT3によって、音声認識に適した音声フレームデータの候補と、各候補に対応する音声認識に適した雑音フレームデータとが関連付けられて制御部8へ通知される。

【0100】

次に、ステップ208において、音声認識部7が最も音声認識に適した音声フレームデータと雑音フレームデータを記憶部4から読み出して音声認識処理を行い、適切な音声認識結果が得られた場合には一連の音声認識処理を終了する。

【0101】

一方、適切な音声認識結果が得られなかった場合には、ステップ212において次の候補の音声フレームデータと雑音フレームデータがあるか否か調べた上で、次の候補の音声フレームデータと雑音フレームデータとを記憶部4から読み出し、再びステップ208からの処理を繰り返す。また、再度の音声認識処理を行

っても適切な音声認識結果が得られなかった場合には、同様に次の候補の音声フレームデータと雑音フレームデータを読み出して音声認識処理を行い、適切な音声認識結果が得られるまでステップ208～212の処理を繰り返す。

【0102】

このように本実施形態によれば、複数個の音声入力用のマイクロフォンM₁～M_nを車室内に配置し、これらのマイクロフォンM₁～M_nで集音されることで生じる音声フレームデータと雑音フレームデータの中から、音声認識に適した音声フレームデータと雑音フレームデータを自動的に抽出して音声認識処理を行うようにしたので、従来の一人の発話者を対象とする音声認識システムと較べて、複数の発話者（搭乗者）に対し極めて優れた利便性を提供することができる。

【0103】

例えば、複数の搭乗者のうちの一人が、あるマイクロフォン（例えばM₁）に対して所望の発話を行った場合に、他のマイクロフォン（M₂～M_N）にもその発話音声が回り込んで集音され、その結果どのマイクロフォンで本当の発話音声が集音されたか解らなくなる状況が一般的には想定されるが、本実施形態によれば、図7及び図8に示した音声条件テーブル6fと雑音選択テーブル6gとを用いて、音声認識に適した音声フレームデータと雑音フレームデータを自動的に抽出して音声認識処理を行うようにしたので、発話をした搭乗者とその搭乗者に近い位置のマイクロフォン（例えばM₁）とが極めて高い確率で対応付けられることになる。

【0104】

このため、本音声認識システムは、音声操作を行おうとする搭乗者を自動的に特定してその発話音声を最適の（近い位置の）マイクロフォンで集音することになるため、音声認識率を向上させることができる。一方、搭乗者の側から見ると、特別な操作をしなくとも、単に発話をするだけで自分の発話指令が適切なマイクロフォンを通して受け入れられることになるため、極めて利便性の高い音声認識システムの構築が可能となっている。

【0105】

更にまた、音声操作を行おうとしていない1又は複数の搭乗者が会話等を行っ

ている状況で、ある一人が音声操作のための発話をを行うような場合でも、図7及び図8に示した音声条件テーブル6fと雑音選択テーブル6gとを用いて、音声認識に適した音声フレームデータと雑音フレームデータを自動的に抽出して音声認識処理を行うことで、音声操作を行おうとしている搭乗者の会話等は雑音として判定されることになって除外される。このため、周囲の会話等の影響を受けない音声認識システムを実現することができ、ひいては、極めて利便性の高い音声認識システムの構築が可能となっている。

【0106】

尚、本実施形態は、図1に示したように、発話スイッチ9を1個だけ設け、例えば運転手がこの発話スイッチ9をオン操作するようにしたが、本発明は、かかる構成に限定されるものではない。例えば、図11のブロック図に示すように、複数のマイクロフォンM₁～M_Nのそれぞれに発話スイッチT K₁～T K_Nを設けておき、いずれかの発話スイッチがオン操作されると、制御部3が、そのオン操作された発話スイッチに対応するマイクロフォンに発話入力を行わせ、オン操作されない発話スイッチに対応する残余のマイクロフォンでは車室内の雑音が集音されたと判断するようにしてもよい。

【0107】

かかる構成によると、発話音声を集音したマイクロフォンと、雑音を集音したマイクロホンとを音声認識処理の前に予め特定することが可能となるため、最も音声認識に適した音声データと雑音データを容易に決定するための処理時間を短縮化することができる。

【0108】

また、図1の構成と図11の構成との折衷案の構成としてもよい。すなわち、マイクロフォンM₁～M_Nの数よりも少ない個数の発話スイッチを車室内の適宜の位置に配置し、何れかの発話スイッチがオン操作されると、制御部8がこれを検出して音声認識処理を開始するようにしてもよい。この場合には、各発話スイッチとマイクロフォンM₁～M_Nが完全に一対一に対応付けられないため、図1に示した構成によって音声認識処理が行われることになるが、発話音声を集音したマイクロフォンと雑音を集音したマイクロホンとを音声認識処理の前に予め特定す

ることが可能となるため、音声認識に適した音声フレームデータと雑音フレームデータを決定するための処理時間を短縮化することができる。

【0109】

更に、マイクロフォン $M_1 \sim M_N$ より少ない数の発話スイッチを設けて図1の構成を適用する場合に、各発話スイッチをマイクロフォンの設置範囲として決めておき、それぞれの設置範囲内に属する1又は複数個のマイクロフォンを、何れの発話スイッチがオン操作されたかで予め特定するようにしてもよい。かかる構成によれば、予め特定された1又は複数個の音声フレームデータと雑音フレームデータから音声認識処理に適したもの抽出すればよいので、処理時間の短縮化が可能となる。

【0110】

また、本実施形態では、車載用電子機器に適用する音声認識システムについて説明したが、本発明の音声認識システムは、汎用のマイクロコンピュータシステムや所謂ワードプロセッサ等の電子機器に適用して、発話による文書入力を行うことも可能である。

【0111】

【発明の効果】

以上説明したように本発明によれば、発話者が所望の発話をすると、複数個の音検知手段から出力される各入力信号の中から、音声認識を行うのに適した音声信号と雑音信号が自動的に決定され、その決定された音声信号と雑音信号に基づいて音声認識が行われるようにしたので、発話者にとって特定の音声入力手段に対して発話するという意識を持つことなく適切な音声認識がなされることになり、利便性の高い音声認識システムを提供することができる。

【図面の簡単な説明】

【図1】

本実施形態に係る音声認識システムの構成を示すブロック図である。

【図2】

マイクロフォンの配置態様の一例を示す説明図である。

【図3】

マイクロフォンの他の配置態様の一例を示す説明図である。

【図4】

マルチプレクサとデマルチプレクサ及び記憶部の構成を示すブロック図である

【図5】

入力信号をサンプリングして記憶部に記憶させる際のタイミングを説明するためのタイミングチャートである。

【図6】

平均音声パワーと平均雑音パワー及び平均S/N値の演算方法を説明するための説明図である。

【図7】

音声条件テーブルの構成を示す説明図である。

【図8】

雑音選択レーブルの構成を示す説明図である。

【図9】

本実施形態に係る音声認識システムの動作を説明するためのフローチャートである。

【図10】

本実施形態に係る音声認識システムの動作を更に説明するためのフローチャートである。

【図11】

本実施形態に係る音声認識システムの変形例の構成を示すブロック図である。

【符号の説明】

- 1 … マルチプレクサ
- 2 … A/D変換器
- 3 … デマルチプレクサ
- 4 … 記憶部
- 5 … 発話音声検出部

6 …データ解析部

6 a …最適音声決定部

6 b …雑音決定部

6 c …平均S/N値演算部

6 d …平均音声パワー演算部

6 e …平均雑音パワー演算部

6 f …音声条件テーブル

6 g …雑音選択テーブル

7 …音声認識部

8 …制御部

9, TK₁~TK_N…発話スイッチ

M₁~M_N…マイクロフォン

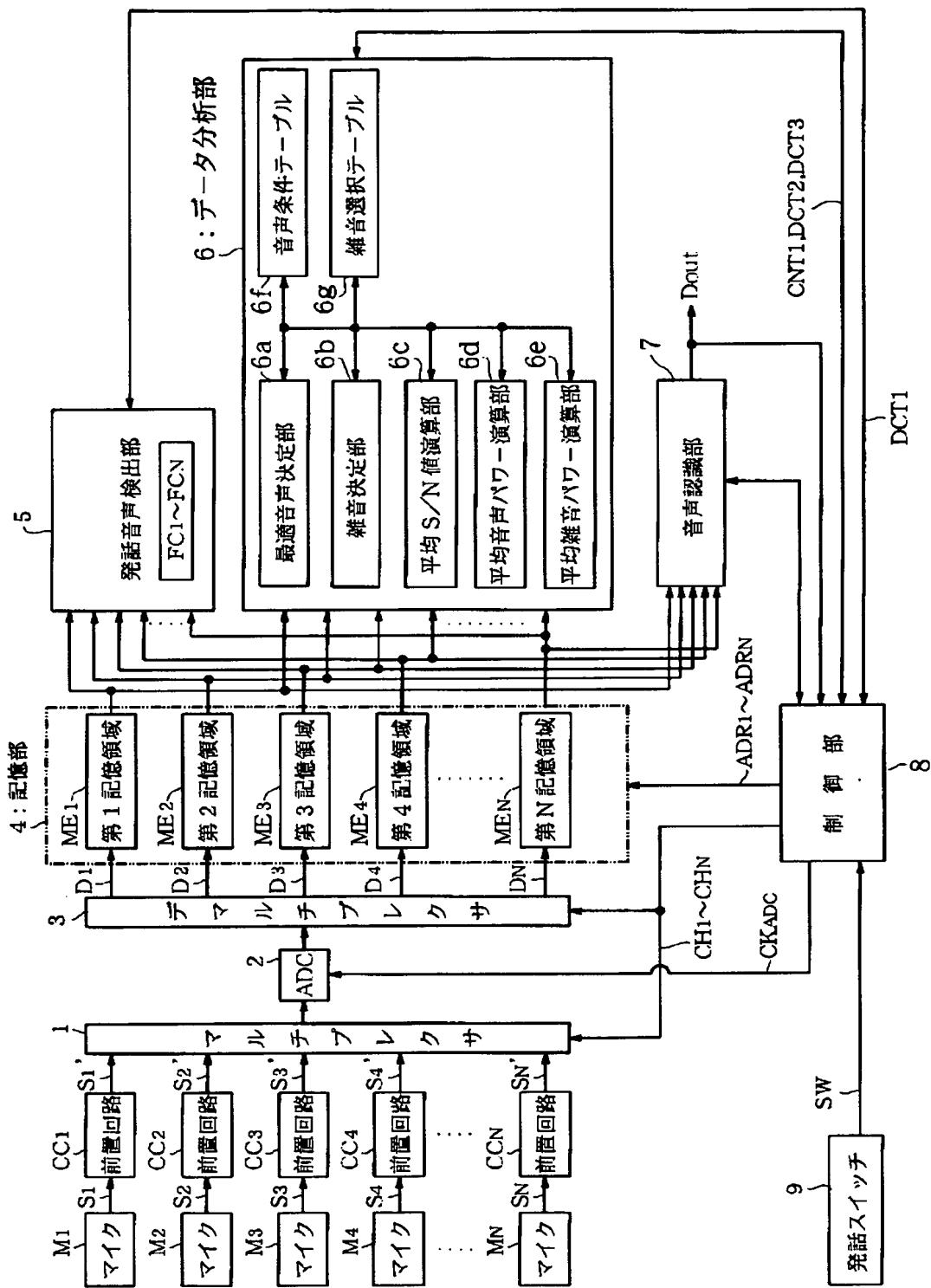
ME₁~ME_N…記憶領域

MF₁~MF_N…フレーム領域

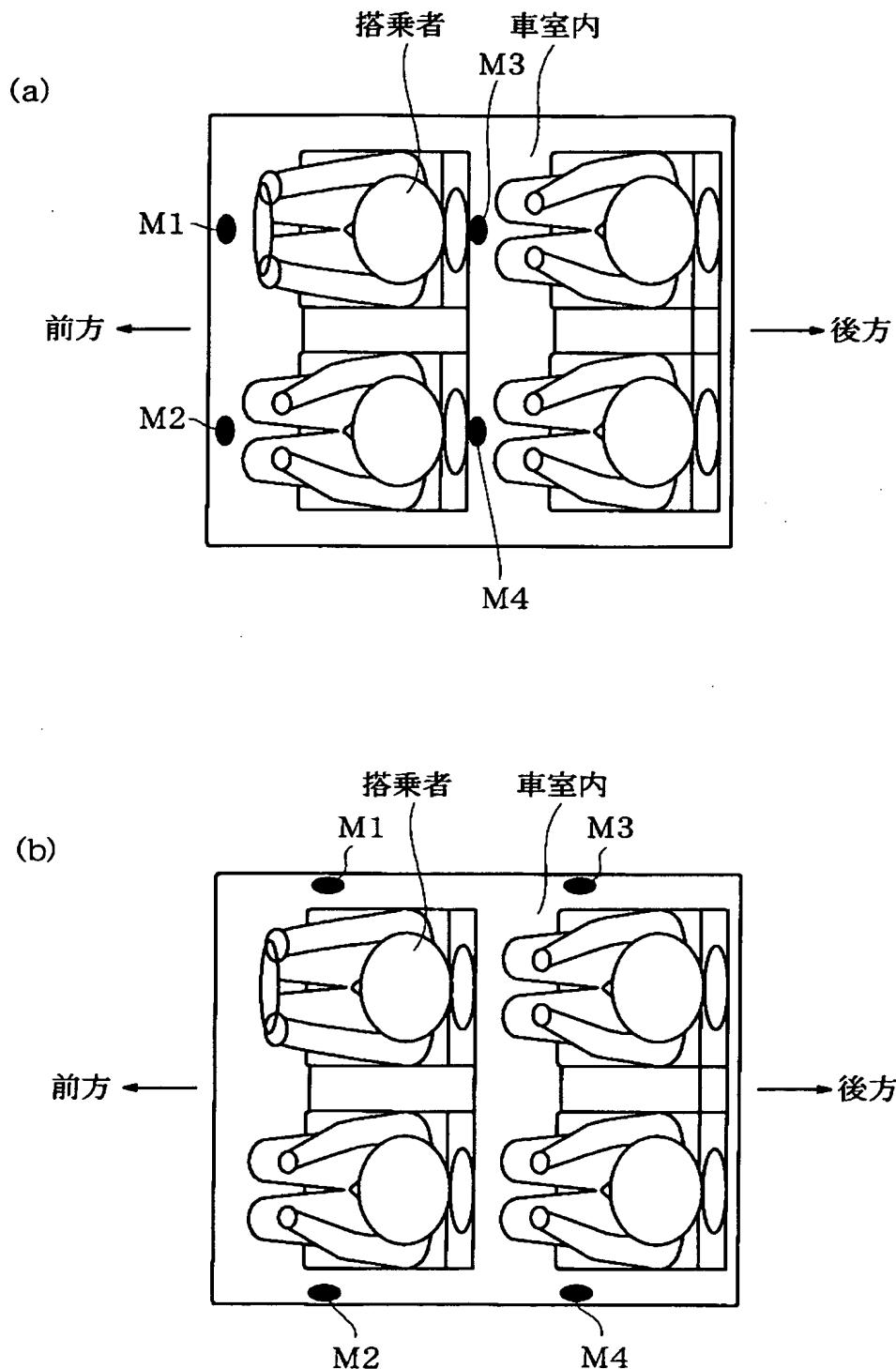
AS₁~AS_N, AW₁~AW_N…アナログスイッチ

【書類名】 図面

【図1】

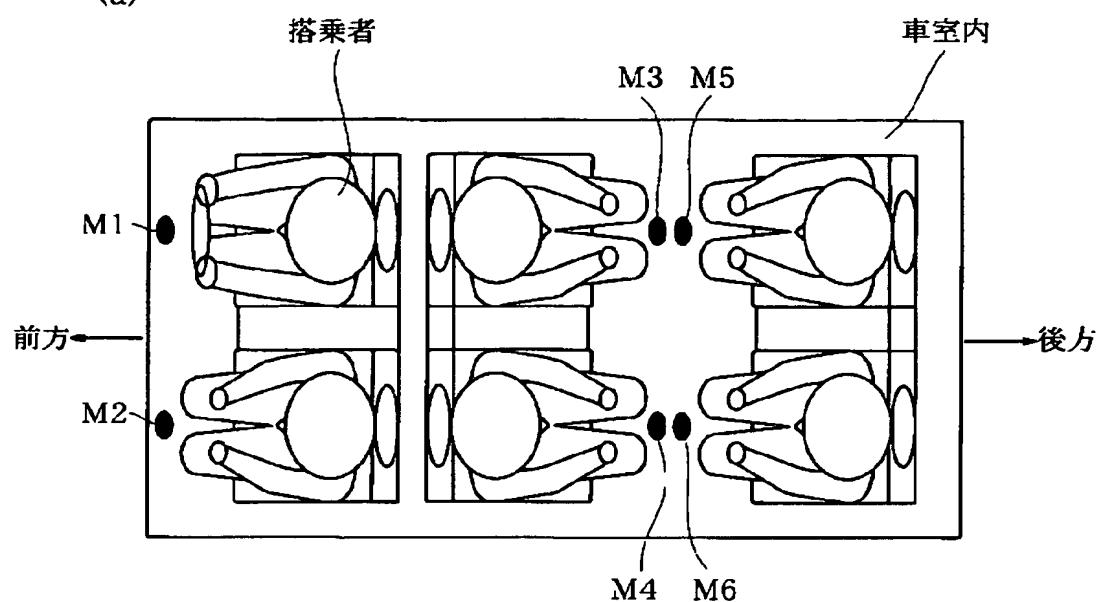


【図2】

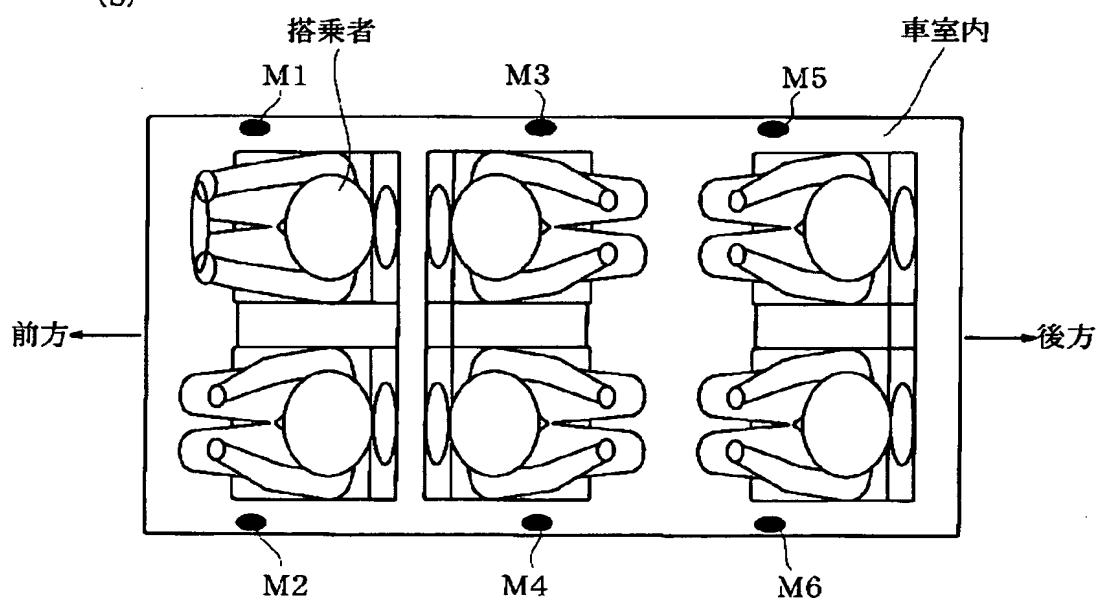


【図3】

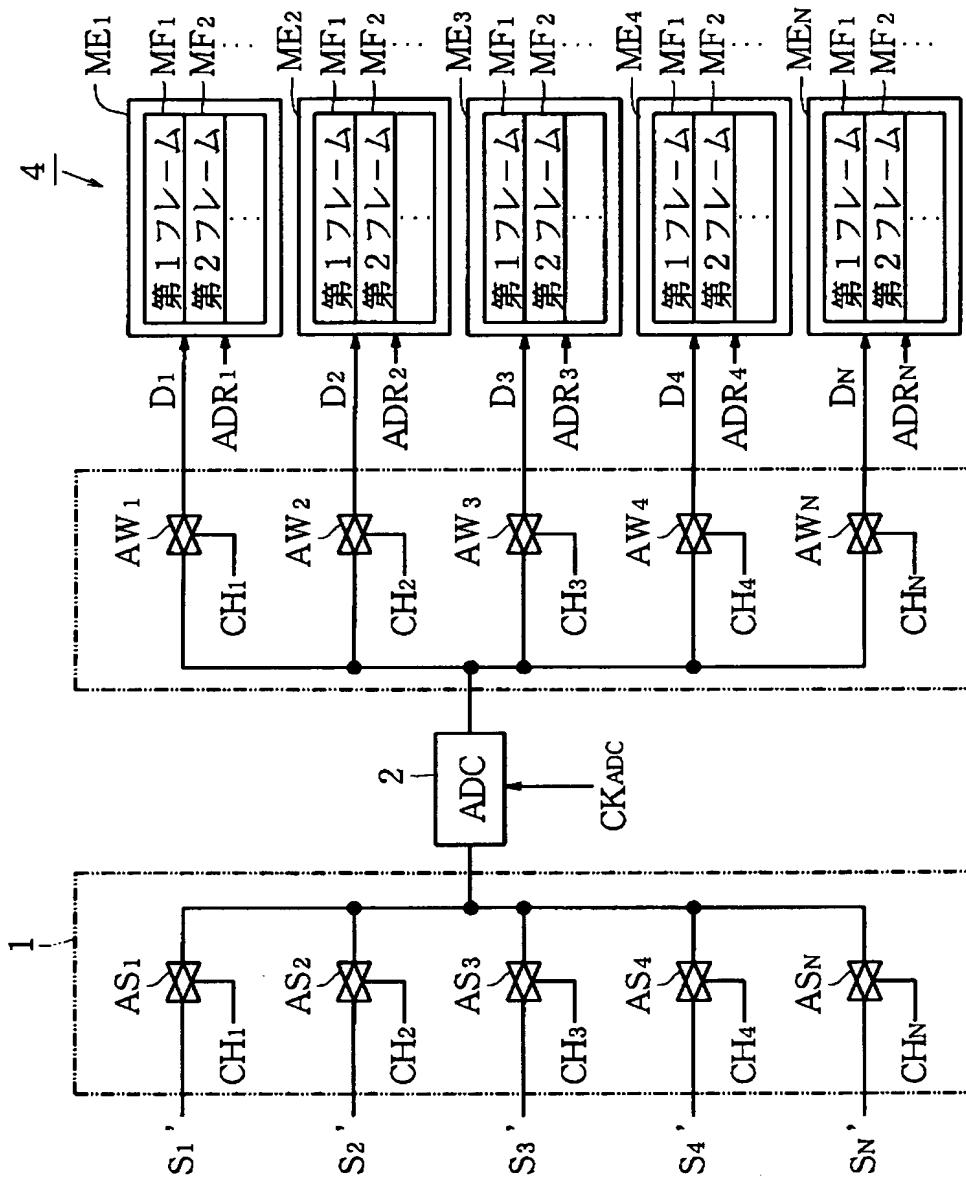
(a)



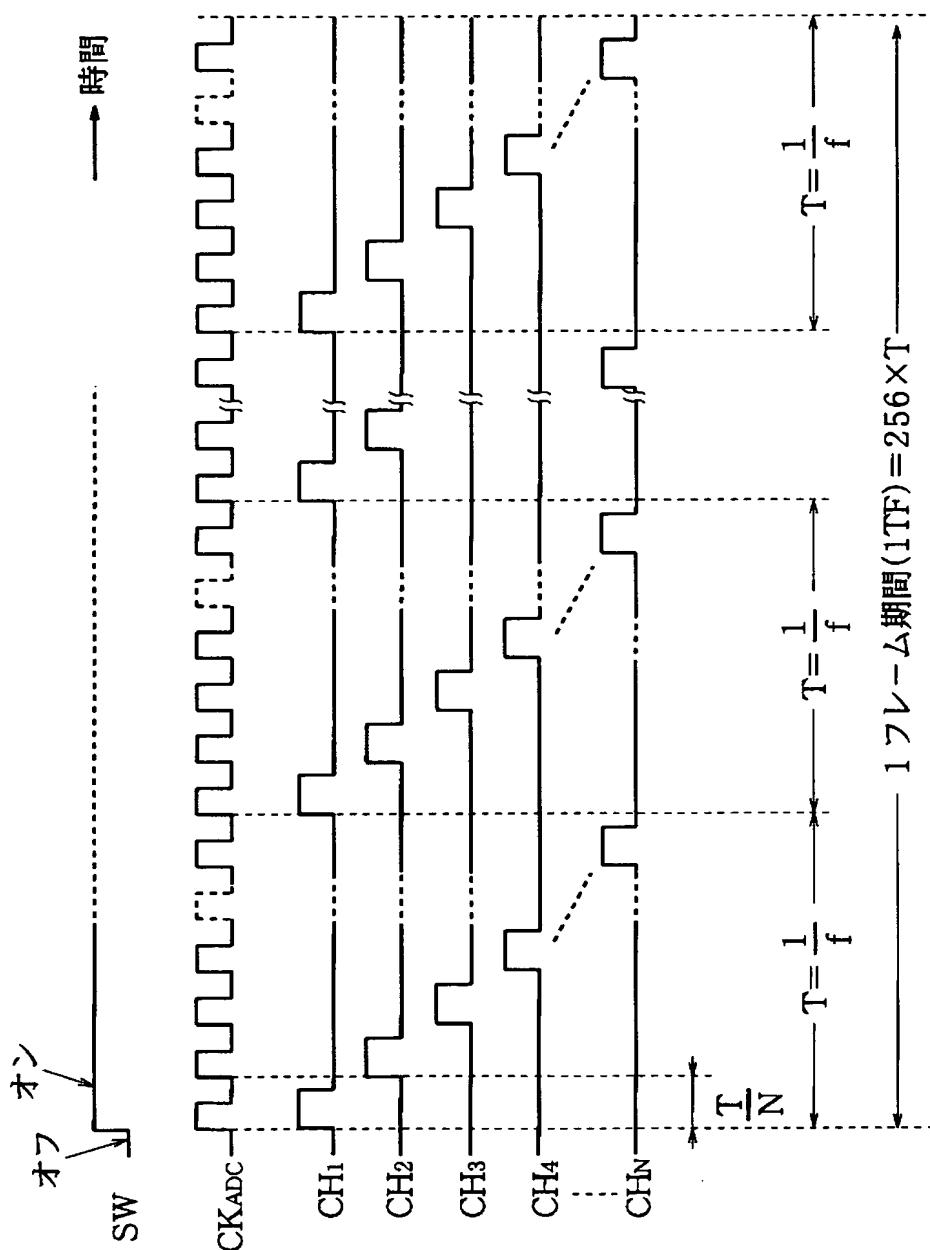
(b)



【図4】

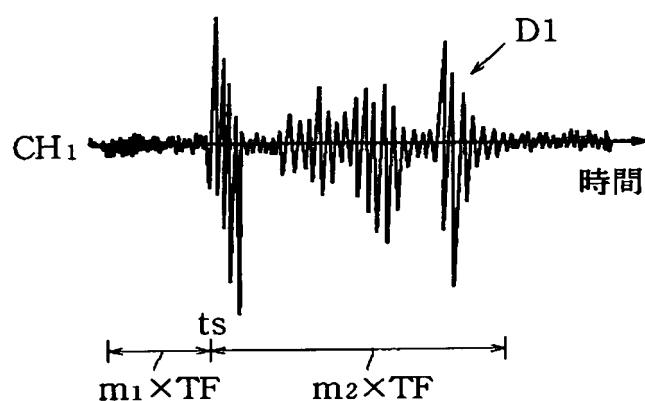


【図 5】

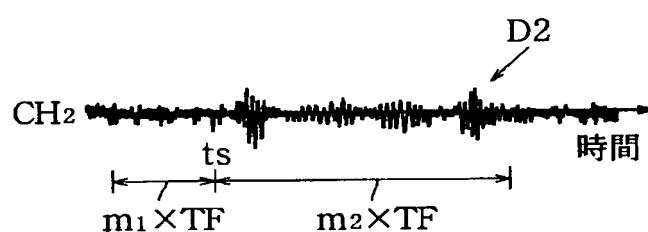


【図 6】

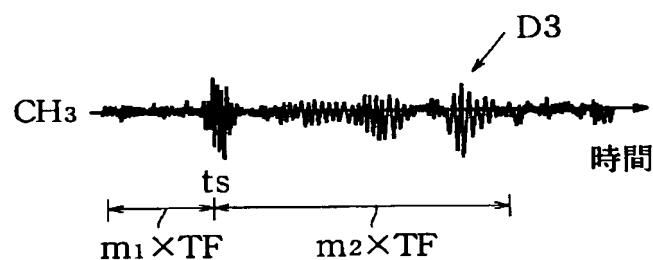
(a)



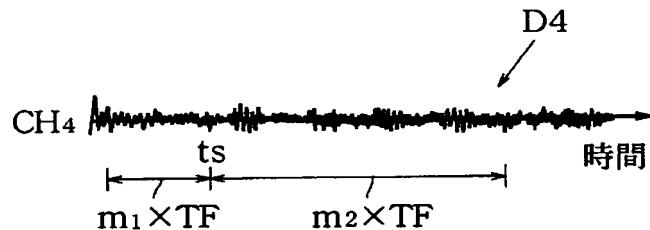
(b)



(c)



(d)



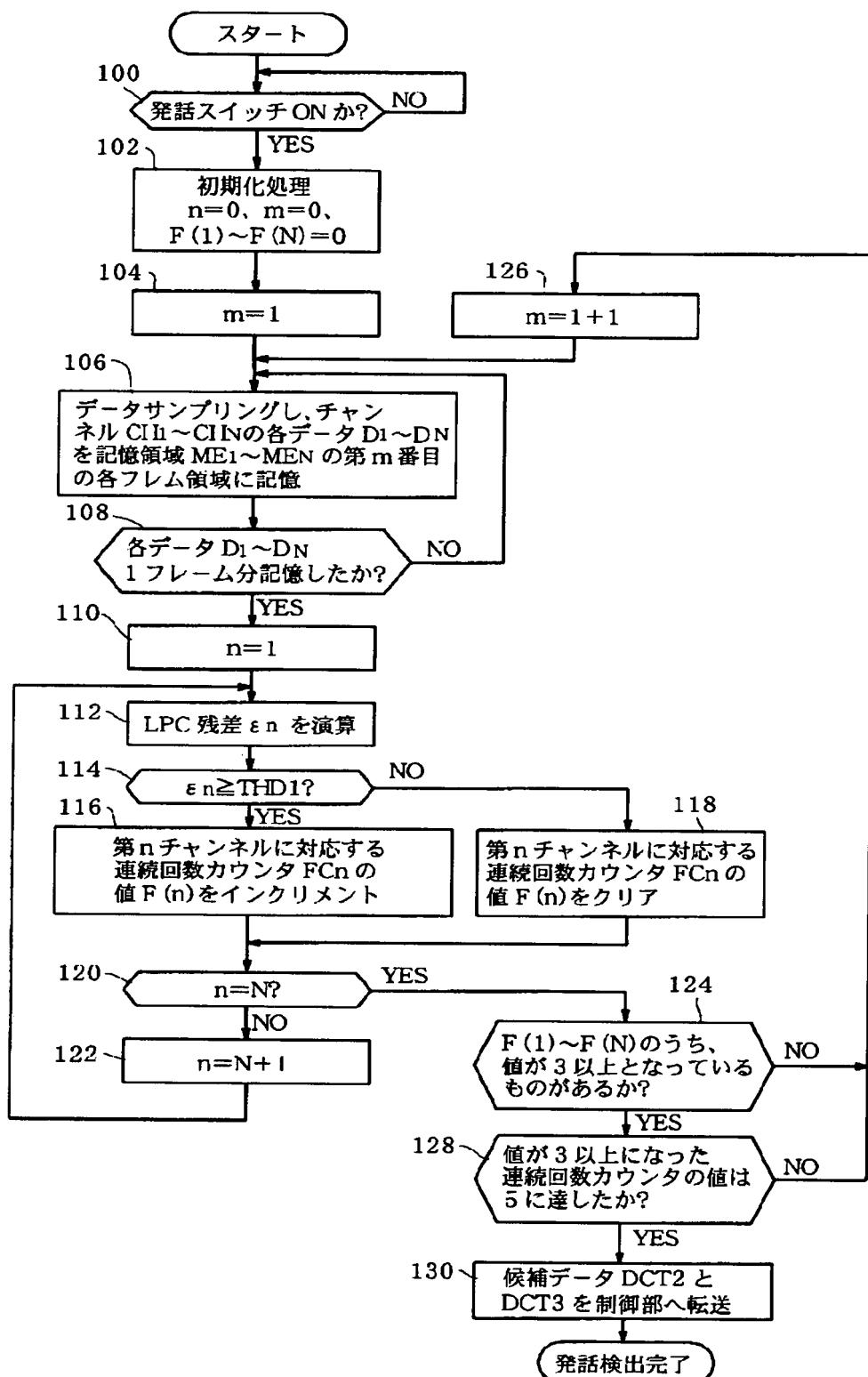
【図 7】

等級	平均 S/N	平均音声パワー
Rnk0	最大値	最大値
Rnk1	THD2 以上	THD3 以上
Rnk2	THD2 以上	THD3 未満
Rnk3	THD2 未満	THD3 以上
Rnk4	THD2 未満	THD3 未満

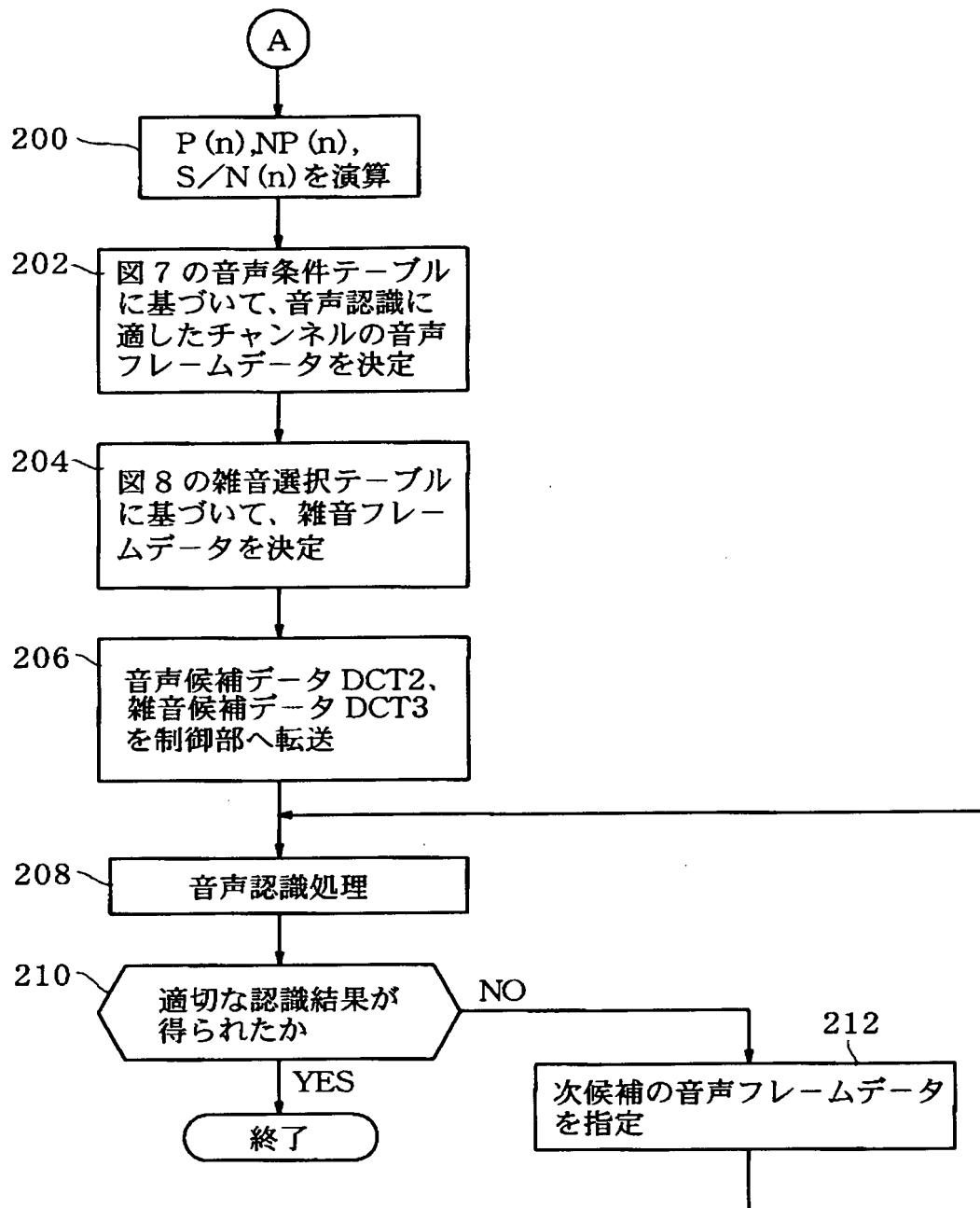
【図 8】

＼	CH1	CH2	CH3	CH4	雑音のチャンネル
場合 1	Rnk0	Rnk1	Rnk2	Rnk1	CH3
場合 2	Rnk0	Rnk1	Rnk1	Rnk2	CH4
場合 3	Rnk0	Rnk1	Rnk1	Rnk3	CH4
場合 4	Rnk0	Rnk1	Rnk1	Rnk4	CH4
場合 5	Rnk0	Rnk1	Rnk2	Rnk1	CH3
場合 6	Rnk0	Rnk1	Rnk2	Rnk2	CH3
場合 7	Rnk0	Rnk1	Rnk2	Rnk3	CH3
場合 8	Rnk0	Rnk1	Rnk2	Rnk4	CH3
場合 9	Rnk0	Rnk1	Rnk3	Rnk1	CH3

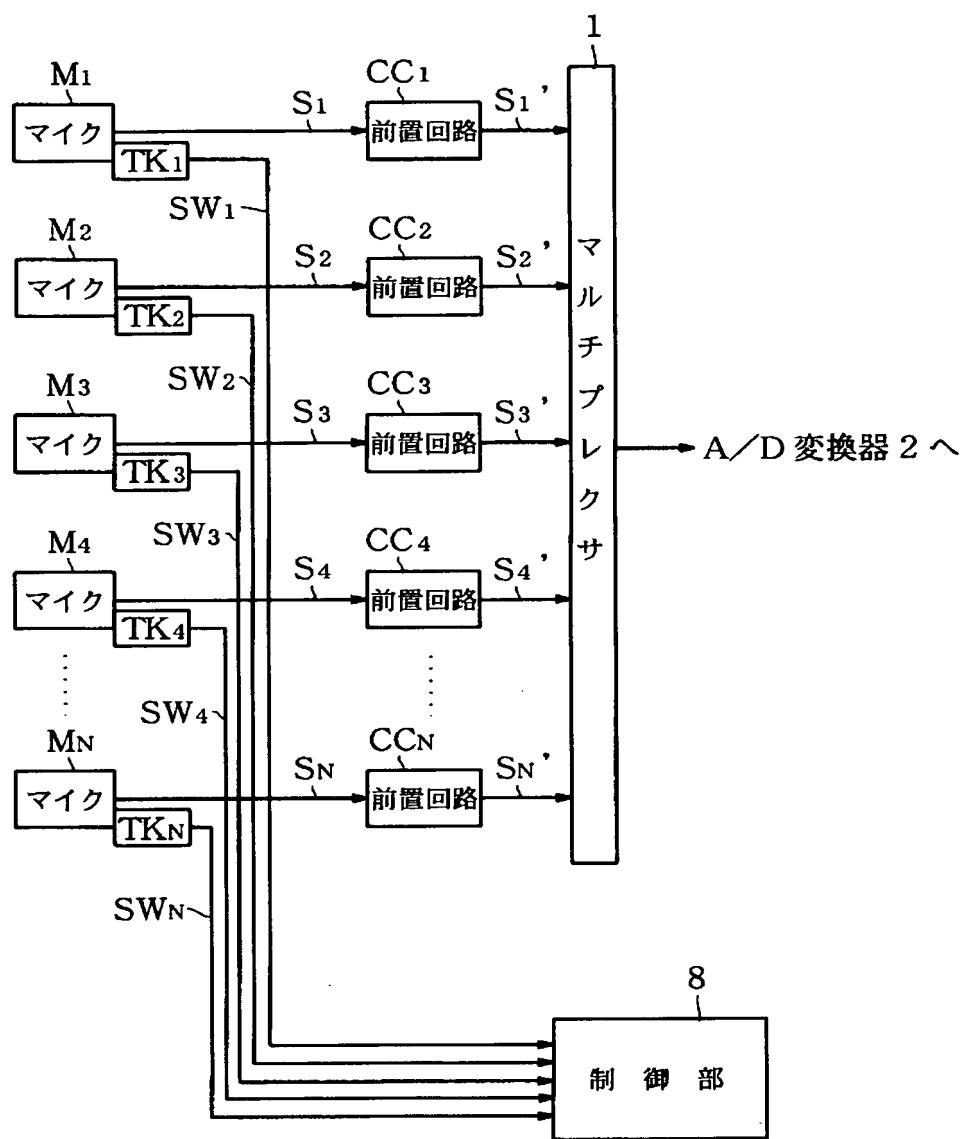
【図9】



【図10】



【図 1 1】



【書類名】 要約書

【要約】

【課題】 利便性の高い音声認識システムを提供する

【解決手段】 複数個のマイクロフォン $M_1 \sim M_N$ を備え、マイクロフォン $M_1 \sim M_N$ から出力される入力信号 $S_1 \sim S_N$ を入力データ $D_1 \sim D_N$ に変換して記憶部4の各記憶領域 $M_E_1 \sim M_E_N$ に所定のフレーム単位で記憶させる。各フレーム単位の入力データ $D_1 \sim D_N$ の各平均S/N値と平均音声パワーとを演算し、平均S/N値と平均音声パワーが所定の閾値より大きくなる入力データを音声認識に適した音声データ、平均音声パワーと平均雑音パワーが最も小さくなる入力データを雑音データと決定し、これらの音声データと雑音データに基づいて音声認識を行う。

【選択図】 図1

出願人履歴情報

識別番号 [000005016]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都目黒区目黒1丁目4番1号
氏 名 パイオニア株式会社